

DSS Statistics Seminar

May 21, 2021, 12:00

<https://uniroma1.zoom.us/j/86881977368?pwd=SWRFcVFjMDZTa0lXZk05TE1zNm5adz09>

Passcode: 432940

Robust clustering with cellwise outliers

Luis Angel García Escudero

Dpto. de Estadística e IO, Universidad de Valladolid

It is well known that a few atypical observations can be quite harmful when applying Cluster Analysis techniques and this justifies the use of robust clustering methods. In this talk, we will present a robust cluster analysis methodology based on cellwise trimming as an extension to a robust version of Principal Component Analysis. This approach is particularly interesting for large or moderately sized problems because it avoids the important loss of information that occurs when trimming entire observations with few outlying cells. Robust fitting of approximating subspaces, which serve to describe the structure of variability within each cluster, is very useful in the automated detection of outlying cells, and it seems logical that the determination of these subspaces and outliers be performed in a unified way, because the two problems are clearly interrelated. We propose an algorithm based on alternating weighted least squares and Least Trimmed Squares regressions to apply this approach. This algorithm is particularized to the case of functional cluster analysis. We provide simulations and applications using real data sets, in demography and meteorology, to illustrate the proposed methodology.



SAPIENZA
UNIVERSITÀ DI ROMA